Synthesis of generalized neural elements using approximation method

(Şynteza uogólnionych elementów neuronowych z użyciem metody aproksymującej)

Sc.D., Prof. FEDIR GECHE¹, Ph.D., Assoc. Prof. ANATOLIY BATYUK², VLADYSLAV KOTSOVSKY¹, PhD KONRAD GROMASZEK²

¹ Uzhgorod National University, ² Lviv National Polytechnic University, ³ Lublin University of Technology

The increase of the functional possibilities of neural elements is important in the synthesis of neural network schemes that are used to implement complex mappings that arise in various problems of classification, form/pattern recognition, time lines prognosis, etc.

This article introduces the concept of generalized neural element with the threshold activation function. At the input neural elements with the generalized threshold activation function can be given as generator groups of characters of groups that are the domain of the Boolean functions, as well as any subset of its elements. Depending on the chosen system of characters on which activation function of neural elements are created, we receive different classes of Boolean functions, elements of which are implemented by one neural element.

In relation to the selected system of characters a functional is created, and it is convex. The task of synthesis of generalized neural elements using approximation method is to minimize this functional. If at minimum point at a given order of approximations this functional is set to 0 and thus every element of the group on which the Boolean function is defined, a weighted sum at the input of neural elements is not zero, then the minimum point is can be selected by the structure vector of neural element which implements the given function.

Mathematical Model of Generalized Neural Element and Its Synthesis

Let \( H = \{ -1, 1 \} \) be a cyclic group of 2\(^n\) order, \( G = H \otimes \ldots \otimes H \) - direct product of cyclic groups \( H \), and \( \chi(G) \) - a group of characters \([1, 2]\) of the group \( G \) over the field of real numbers. On the set \( R \setminus \{ 0 \} \) we define the function:

\[
R \text{sign} x = \begin{cases} 
1, & \text{if } x > 0, \\
-1, & \text{if } x < 0.
\end{cases}
\]

(1)

Let \( i \in \{ 0, 1, 2, \ldots, 2^n - 1 \} \) and \((i_0, i_1, \ldots, i_{2^n-1})\) be its binary code, i.e. \( i = 2^{i_0} + 2^{i_1} + \ldots + 2^{i_{2^n-1}} \), \( i_0, i_1, \ldots, i_{2^n-1} \in \{ 0, 1 \} \). The value of the character \( \chi_i \) in the element \( g = ((-1)^{i_0}, \ldots, (-1)^{i_{2^n-1}}) \in G \), \((i_0, \ldots, i_{2^n-1}) \in \mathbb{Z}_2^n \) is defined as:

\[
\chi_i(g) = (-1)^{i_0, \ldots, i_{2^n-1}} g_{i_0} \cdots g_{i_{2^n-1}}.
\]

(2)

Consider the 2\(^n\) - dimensional vector range \( V = \{ g; \mathfrak{g}: G \to R \} \) over the field \( R \). Elements \( \chi_i(i = 0, 1, 2, \ldots, 2^n - 1) \) of the group \( \chi(G) \) form an orthogonal basis of the range \( V \). [3, 5]. Boolean function in the alphabet \( \{ 1, 2 \ldots, 2^n - 1 \} \) gets unambiguous mapping \( f: G \to H \), meaning \( f \in V \). Thus, any arbitrary Boolean function \( f \in V \) can be uniquely written in the form:

\[
f(g) = s_0 \chi_0(g) + s_1 \chi_1(g) + \ldots + s_{2^n-1} \chi_{2^n-1}(g).
\]

(3)

Vector \( s = (s_0, s_1, \ldots, s_{2^n-1}) \) is called the spectrum of Boolean function in the system of characters \( \chi(G) \) (in the system of basic functions of Walsh-Hadamard).

Using different characters \( \chi(G) \), except the main one, construct \( m \)-element set \( \{ \chi_1, \ldots, \chi_m \} \) and in relation to the chosen system of characters lets examine the following mathematical model of neural element:

\[
f(x_1(g), \ldots, x_n(g)) = R \text{sign} (\sum_{j=1}^n a_j \chi_j(g) + a_k),
\]

(4)

where vector \( w = (a_0, \ldots, a_{2^n-1}; a_k) \) is called the structure vector of neural element and \( g \in G \).

May \( w(g) = a_0 \chi_0(g) + \ldots + a_{2^n-1} \chi_{2^n-1}(g) + a_k \). If \( w = (a_0, \ldots, a_{2^n-1}; a_k) \) is a structure vector of neural element relatively to the system of characters \( \{ \chi_1, \ldots, \chi_m \} \) of group \( G \) over \( R \), that realizes the Boolean function \( f: G \to H \), then out of (1) and (4) immediately implies that

\[
\forall g \in G, w(g) \neq 0.
\]

(5)

Now, following above mentioned, we will consider only such neural elements, vector structures of which satisfy the condition (5). The set of all \( m+1 \)-dimensional real vectors, that satisfy the condition (5), lets mark as \( W_m = W_{m+1}(\chi_1, \ldots, \chi_m) \).

It is obvious that the neural elements in relation to the system of characters \( \{ \chi_1, \chi_2, \ldots, \chi_{m+1} \} \) coincides with the threshold element [4-7].

Similarly as in the threshold case, a question also arises whether a given Boolean function \( f(x_1, \ldots, x_n) \) realizes by one neural element relatively to the chosen system of characters \( \{ \chi_1, \ldots, \chi_m \} \), and if does, then how to find the corresponding structure vector of neural element?

If from the spectral coefficients \( s_1, \ldots, s_{2^n-1} \) of Boolean function \( f \) that correspond to the system of characters \( X = \{ \chi_1, \ldots, \chi_m \} \) build a vector \( s(X) = (s_1, \ldots, s_{2^n-1}) \), which will be named as the characteristic vector of Boolean function \( f \) respectively to the system \( X \), then [8] provides the criteria for the feasibility of a Boolean function by one generalized neural element, namely:

Theorem 1

Boolean function \( f: G \to H \), is realised by one neural element relatively to the system of character \( X = \{ \chi_1, \ldots, \chi_m \} \subseteq X(G) \) with the structure vector \( w \in W_{m+1} \) whereas and only if its characteristic \( s(X) \) satisfies the condition

\[
(w, s(X)) = \sum_{g \in G} w(g) s(X)(g),
\]

(6)

where \( (w, s(X)) \) - is a scalar product of vectors \( w \) and \( s(X) \).

It also shows that if neural elements of the vector with respect to the structure of it also shows that if neural element within the structure vector \( w \in W_m \) relatively the system of characters \( \{ \chi_1, \ldots, \chi_m \} \) realizes the Boolean function \( f \), then for the Boolean function \( h: G \to H \) \((h \neq f)\) performs the inequality

\[
(w, s(X)) > (w, s(X)).
\]

(7)

The main difficulty in the development of the practical accepted methods of the synthesis of neural elements relatively to the \( X \) is in the non-linearity of expression \( \sum_{g \in G} |w(g)| \). In a method of minimizing the functional [2] this expression is approximated by a polynomial of appropriate degree and has vector \( w \), that minimizes the built functional. The vector \( w \) is taken as an approximate structure vector of neural element relatively the approximation of these order.
Functional \( U(\mathbf{w}, s, f(X)) \), subject to minimization is defined as following:

\[
U(\mathbf{w}, s, f(X)) = \sum_{g \in G_n} |w(g)| - \langle \mathbf{w}, s, f(X) \rangle,
\]

(8)

where \( \mathbf{w} = (\omega_1, \ldots, \omega_m; \omega_0) \) is an arbitrary vector of dimensions \( m+1 \)-dimensional space \( R^{m+1}, s_j(X) = (s_1, \ldots, s_m; s_0) \).

Functional \( U(\mathbf{w}, s, f(X)) \) is convex on the set of all \( m+1 \)-dimensional real vectors. Indeed, if \( \mathbf{w}_1, \mathbf{w}_2 \in R^{m+1} \) and \( \lambda \mathbf{w}_1 + (1-\lambda) \mathbf{w}_2, (\lambda \in [0,1]) \), then

\[
U(\lambda \mathbf{w}_1 + (1-\lambda) \mathbf{w}_2, s, f(X)) = \sum_{g \in G_n} |(\lambda \mathbf{w}_1 + (1-\lambda) \mathbf{w}_2)(g)| - \langle \lambda \mathbf{w}_1 + (1-\lambda) \mathbf{w}_2, s, f(X) \rangle \\
\leq \sum_{g \in G_n} |\mathbf{w}_1(g)| + (1-\lambda) \sum_{g \in G_n} |\mathbf{w}_2(g)| - \lambda \langle \mathbf{w}_1, s, f(X) \rangle - (1-\lambda) \langle \mathbf{w}_2, s, f(X) \rangle \\
= \lambda U(\mathbf{w}_1, s, f(X)) + (1-\lambda) U(\mathbf{w}_2, s, f(X)).
\]

If \( \mathbf{w} \in W_{m+1} \), then from the Theorem 1 and inequality (7) follows that \( \lambda U(\mathbf{w}, s, f(X)) \geq 0 \) and sing "w" is there only if \( \mathbf{w} \) is a structure vector of neural element relatively to the system \( X \), that realizes function \( f \).

Let consider the minimization of the functional \( U \) under condition that its non-linear polynomial expression \( \sum_{g \in G_n} |w(g)| \) is replaced by the polynomial of appropriate degree, depending on the order of approximation.

May

\[
g | w(g) = \zeta_0 g^{-2} \mathbf{w}(g) + \zeta_1 g^{-1} \mathbf{w}^2(g) + \ldots + \zeta_k g^{-k} \mathbf{w}^k(g),
\]

(9)

where \( k \in \{1,2,3,\ldots\} \) and \( g \) - a normalizing multiplier, that

\[
\forall g \in G_n, \quad 0 < g | w(g) \leq 1.
\]

The values \( \zeta_k \) in (9) choose so, as to minimize the standard deviation between the derivative \( w(g) \) function \( g | w(g) \) and its polynomial approximation.

If \( k = 1 \), then we will talk about the approximation of 1st order, if \( k = 2 \), then about the approximation of 2nd order etc.

Let's consider the approximation of the 1st order. In this case:

\[
|w(g)| = \zeta_0 g^{-2} \mathbf{w}(g), \quad (g \in G_n).
\]

Summing the left and right parts of the last expression according to all \( g \in G_n \), we get:

\[
\sum_{g \in G_n} |w(g)| = \zeta_0 g \sum_{g \in G_n} \mathbf{w}(g),
\]

(11)

where

\[
\mathbf{w}(g) = \sum_{j=0}^{m} \omega_j \chi_j(g) + \omega_0,
\]

and \( \chi_0 = \chi_0 \) - main character of group \( G_n \). Placing the meaning \( \mathbf{w}(g) \) in (11) and considering orthogonal correlation of the characters [1], we get:

\[
\sum_{g \in G_n} |w(g)| = \zeta_0 g \sum_{g \in G_n} \sum_{j=0}^{m} \omega_j \chi_j(g) \chi_j(g) = \zeta_0 g \sum_{j=0}^{m} \omega_j = \zeta_0 g \sum_{j=0}^{m} \omega_j = \zeta_0 g \sum_{j=0}^{m} \omega_j.
\]

Then functional \( U(\mathbf{w}, s, f(X)) \) may be written as:

\[
U(\mathbf{w}, s, f(X)) = \sum_{g \in G_n} |w(g)| - \langle \mathbf{w}, s, f(X) \rangle,
\]

(12)

where \( i_0 = 0 \).

Functional \( U(\mathbf{w}, s, f(X)) \), according to (12), is continuous relatively to \( \omega_i(i = 0,1,\ldots,m) \). So, functional \( U(\mathbf{w}, s, f(X)) \) can be minimized by equating to zero the partial derivatives \( U \) for \( \omega_j \):

\[
\frac{\partial U(\mathbf{w}, s, f(X))}{\partial \omega_j} = 0, \quad j = 0,1,\ldots,m.
\]

Herewith

\[
\zeta_0 g \omega_i = 0, \quad i = 0,1,\ldots,m.
\]

If \( \omega \) - nonzero vector, then from the |w(g)| = \zeta_0 g \mathbf{w}(g) follows that \( \zeta_0 > 0 \). Then for the coordinates \( \omega_i \) of the vector \( \omega \) can be selected corresponding spectral coefficients \( s_j \) of the Boolean function \( f \), because the neural elements of the structure vectors \( \mathbf{w} \) and \( \lambda \mathbf{w} \) relatively the system \( X = \{X_1, \ldots, X_n\} \) realize one and the same Boolean function, if \( \mathbf{w} \in W_{m+1} \) and \( \lambda > 0 \). Thus, \( \omega_j = s_j, (j = 0,1,\ldots,m) \) and as a result of the 1st order approximation for the structure vector of neural elements the vector \( \mathbf{w} = (\omega_1, \ldots, \omega_m, \omega_0) \) is not selected. Let's consider two more of the possible cases:

1) \( \mathbf{w} \in W_{m+1} \),
2) \( \mathbf{w} \in \Omega \).

In the first case, if \( \forall g \in G_n, \quad f(g) = R\text{sign} \mathbf{w}(g) \), the neural element with the structure vector \( \mathbf{w} \) realizes Boolean function \( f \) and the task of synthesis is finished. If there is such element as \( g \in G_n \), that \( f(g) = R\text{sign} \mathbf{w}(g) \), then we have an option to finish the synthesis process of the neural element using the approximation method of the 1st order (not constructing it), or we switch to the approximation of the 2nd order.

In the second case we construct two vectors \( \mathbf{w}_0 \) and \( \mathbf{w}_1 \), which belong to the set \( W_{m+1} \), and are arbitrary close to the vector \( \mathbf{w} \), i.e., for any infinitesimal \( \varepsilon > 0 \) distance \( \rho(\mathbf{w}, \mathbf{w}_0) < \varepsilon \) and \( \rho(\mathbf{w}, \mathbf{w}_1) < \varepsilon \). While constructing the vectors \( \mathbf{w}_0, \mathbf{w}_1 \in W_{m+1} \) we impose on it following conditions: if \( g \in G_n \), such that \( w(g) \neq 0 \), then

\[
R\text{sign} \mathbf{w}(g) = R\text{sign} \mathbf{w}_0(g) = R\text{sign} \mathbf{w}_1(g),
\]

(14)

May

\[
\alpha_{min} = \min \{ |w(g)|, g \in G_n, w(g) > 0 \},
\]

\[
\alpha_{max} = \max \{ |w(g)|, g \in G_n, w(g) < 0 \},
\]

and

\[
\alpha_{min} = \alpha_0 + 0 \omega, \quad \alpha_{max} = \alpha_0 - 0 \omega,
\]

where \( \omega = \min \{ \alpha_{min}, \alpha_{max} \} \). Then vectors \( \mathbf{w}_0 = (\alpha_0, \ldots, \alpha_0, \omega, \ldots, \omega) \) and \( \mathbf{w}_1 = (\alpha_0 + \omega, \ldots, \alpha_0, \omega, \ldots, \omega) \) satisfy the conditions (14), (15) and at

\[
p \rightarrow \infty, \quad \rho(\mathbf{w}, \mathbf{w}_0) \rightarrow 0.
\]

If \( \forall g \in G_n \) fulfills one of the conditions

\[
f(g) = R\text{sign} \mathbf{w}_0(g),
\]

(16)

or

\[
f(g) = R\text{sign} \mathbf{w}_1(g),
\]

(17)

then for the structure vector of the neural element, that realizes Boolean function \( f \), is chosen vector \( \mathbf{w}_0 \) or depending on whether there is a correlation (16 or 17). If none of these conditions are met, then synthesis of the neural element using approximation...
Synthesis of the Neural Elements Relatively to $X$ Using Approximation Method of the 1st Order
Synteza elementów neuronowych względem $X$ z wykorzystaniem metody aproksymacji pierwszego rzędu

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
<th>$x_6$</th>
<th>$f$</th>
<th>$s_i$</th>
<th>$w'(g)$</th>
<th>$w(g)$</th>
<th>$w'(g)$</th>
<th>$w(g)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-4</td>
<td>8</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>12</td>
<td>-8</td>
<td>-4</td>
<td>-12</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>-16</td>
<td>-12</td>
<td>-20</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-4</td>
<td>16</td>
<td>20</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-4</td>
<td>-16</td>
<td>-12</td>
<td>-20</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-24</td>
<td>-20</td>
<td>-28</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>4</td>
<td>-8</td>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>4</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-24</td>
<td>-20</td>
<td>-28</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>-16</td>
<td>-12</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-4</td>
<td>16</td>
<td>20</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-4</td>
<td>-16</td>
<td>-12</td>
<td>-20</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-4</td>
<td>-16</td>
<td>-12</td>
<td>-20</td>
<td></td>
</tr>
</tbody>
</table>

Based on Table we conclude that

$$\forall g \in G_L \quad f(g) = R \text{sign} w'(g),$$

meaning that function $f(x_1, x_2, x_3, x_4)$ is realized by one neural element related to the system $X = \{x_1, x_2, x_3\}$ with the structure vector $w = (12, -4, 4, -8)$ or with the structure vector $w = (3, -1, 1, 1, -2)$.

Conclusion

A synthesis of generalized neural elements using approximation method of the 1st order was developed. With this method it was shown that above mentioned function $f(x_1, x_2, x_3, x_4)$ is realized by one generalized neural element relatively to the system $X = \{x_1, x_2, x_3\}$, while it is realized by one regular neural element with the threshold function of activation relatively to the system $(x_1, x_2, x_3, x_4)$. This means that generalized neural elements may be successfully used in implementing complex mappings and by increasing the functionality of neural elements we may get fewer neurons in the neural network schemes, which are used to solve current applied problems of form/pattern recognition, time lines prognostication etc.

Results of the approximation method of the 1st order may be successfully used in the iterative methods of the synthesis of generalized neural elements, namely it is that as the first vector (at the beginning of iteration) we choose vector that was found by the above mentioned method.

References